

University of Groningen

Ultimate attainment at the interfaces in second language acquisition

Hopp, H.C.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2007

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Hopp, H. C. (2007). *Ultimate attainment at the interfaces in second language acquisition: grammar and processing*. [Thesis fully internal (DIV), Rijksuniversiteit Groningen]. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

2. Ultimate Attainment and Second Language Processing

Introduction

In Chapter 1, I reviewed how differences in *grammatical representation* between non-natives and natives potentially restrict convergence on the TL in L2 acquisition. This chapter discusses psycholinguistic evidence on how differences in *language processing* between non-natives and natives may limit ultimate attainment in the L2. Starting with Penfield & Roberts (1959) and Lenneberg (1967), a core tenet of the Critical Period Hypothesis formulated from a neurophysiological perspective has been that language acquisition after the offset of a critical period implicates the engagement of a different set of language processing mechanisms and neural regions compared to language acquisition within the critical period.

In a recent comprehensive overview of grammatical processing in native and non-native acquisition, Clahsen & Felser (2006b) outline four potential differences between native and non-native processing (1) that might constrain convergence at L2 ultimate attainment.

- (1) a. L2 processing is fundamentally different from native processing
L2 acquirers recruit different processing mechanisms and subservient neural regions compared to native speakers.
- b. L2 processing only makes partial use of native processing mechanisms
L2 acquirers do not recruit the full set of processing mechanisms or neural substrates available to native speakers.
- c. L2 processing is affected by L1 transfer
The interference of L1 processing mechanisms entails differences between non-native and native processing.
- d. L2 processing is computationally less efficient than native processing
L2 acquirers cannot process the TL as quickly and as automatically as native speakers.

In this chapter, I review data and models of L2 processing in the context of these four potential explanations. The first two potential differences in (1) may be associated with critical period effects, whereas the final two potential differences do not or do not necessarily implicate critical period effects. Given the focus of this thesis on syntax and its interfaces, I will be exclusively concerned with sentence comprehension or, more particularly, parsing that describes the process of assembling words to syntactic structure.

This chapter tackles L2 processing in six steps. I start out with a brief general discussion of the relation between grammar and parsing, and I introduce the psycholinguistic methods and measures employed in (L2) processing research in Section

2.1. In a second step, I review research on maturational constraints affecting the neurophysiological bases of language processing (Section 2.2). The third section of this chapter discusses the processing mechanisms recruited in L2 processing in syntax and argument structure. In the fourth step, I consider potential factors affecting L2 processing, such as L1 effects, proficiency, automaticity and working memory. Fifth, I outline and discuss some approaches to L2 processing that have been advanced to account for the patterns of L2 processing and its differences compared to native language processing in Section 2.7. In the concluding Section 2.8, I put the previous studies of L2 grammars (Chapter 1) and L2 processing in perspective, and I outline some desiderata for studying the link between representation and computation at L2 ultimate attainment.

2.1. The interface between grammar and parsing

Research on sentence processing aims to identify the architecture and processes characterizing sentence comprehension and production. One of the guiding questions has been how the mental grammar interfaces with mechanisms of language use. The majority of psycholinguistic models construe the grammar-parser interface in terms of a distinction between mechanisms specialized for the comprehension and production of sentences (the performance systems) and an underlying repository of linguistic knowledge that the performance systems interact with (the competence system). On this view, grammar is the competence system and the parser is the performance system for sentence comprehension (for discussion, see, e.g., Crocker, Pickering & Clifton, 2000; Phillips, 1996).¹

Sentence comprehension involves the left-to-right analysis of linearly ordered input in order to deliver a hierarchically ordered grammatical structure as an output for interpretation. For this task, the parser employs specific parsing principles and mechanisms that are geared to the incremental analysis of potentially ambiguous input, for instance, ambiguity resolution strategies, phrase-boundary heuristics, etc. In addition, the parser draws on grammatical information in building phrase-structure representation of the input. Finally, sentence comprehension is constrained by the cognitive resources available (Figure 2.1).

¹ Here and in the following, I will not be concerned with the production side of language processing, which may require different and additional assumptions and mechanisms.

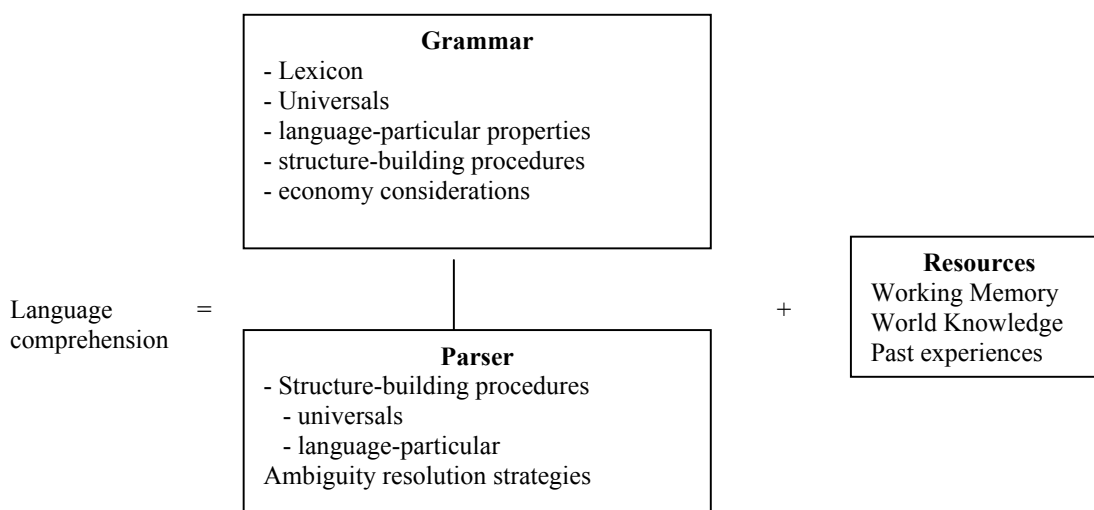


Figure 2.1. *Grammar-parsing interface in language comprehension (adapted from Phillips, 1996: 16).*

Before I review previous studies on L2 processing in the context of the general parsing model sketched above, the following section gives a brief overview of the experimental paradigms used in the study of native and non-native sentence processing.

2.1.1. L2 processing: Methods

In terms of methods, neurophysiological methods and behavioural reaction-time methods can be distinguished. Neuroimaging studies address the question whether the L2 is localized in identical or different brain regions compared to the L1. Behavioural reaction-time and electrophysiological studies explore the temporal dynamics of L2 processing.

Neurophysiological measures

Neuroimaging techniques such as **PET** (Positron Emission Tomography) and **fMRI** (functional magnetic resonance imaging) measure changes in neuronal activity between two processing conditions as indicated by changes in blood flow to particular brain areas. Since changes of hemodynamic states take long compared to the fast pace of speech or reading, the comparatively poor temporal resolution of these methods requires studying fairly extended tasks, e.g. listening to continuous speech or making repeated judgements or producing words or sentences. Nevertheless, these techniques afford the precise identification of the brain areas recruited in processing.

Event-related potentials (ERPs) allow for the identification of the temporal signatures of neurophysiological events in the context of language processing. ERPs offer an extremely high temporal resolution in real-time processing. Time-locked to a specific (linguistic) stimulus, ERPs record voltage changes incurred on a large number of electrodes on the scalp by that stimulus compared to a control condition. ERPs comprise positive and negative voltage peaks (components) that differ in polarity, amplitude, latency and scalp distribution. The literature on the interpretation of ERP components in the context of language processing is abundant and the following remarks are intended to give the briefest of orientation (for general overview, see, e.g., Brown & Hagoort, 1999; for an L2-related overview of ERP studies, see Mueller, 2005).

The most relevant language-related components are, amongst others, (a) a centro-parietal negativity occurring approximately 400ms after stimulus onset that is referred to as N400 and associated with semantic integration processes, (b) a centro-parietal positivity peaking at around 600ms, the P600, which is linked to syntactic integration, syntactic disambiguation or syntactic reanalysis, amongst other things, and (c) a left anterior negativity (LAN) peaking prior to the P600 which has been related to morphosyntactic incongruity. In addition, studies on word category violations elicit a very early left anterior negativity (ELAN) with a latency between 100-300ms. The ELAN has been linked to rapid parsing preferences in syntactic category identification. The different ERP components can thus be interpreted as reflecting distinct processes characterizing semantic and syntactic processing.

Behavioural measures

Speeded grammaticality judgements require participants to make a grammaticality judgement as fast as possible in response to a sentence presented at a rate above normal reading speed. The presentation of a sentence can be broken up into words or phrases. The elevated presentation rate is supposed to prevent the parser from full parsing, such that initial parsing preferences can be tapped. Processing difficulty is analysed as depressed judgement accuracy and increased response time to a sentence compared to a matched control sentence. However, speeded judgements are a rather coarse-grained measure since they only record end-of-sentence effects, such that the locations of incremental processing difficulty cannot be identified.

A more sensitive reading-time measure is the **self-paced reading paradigm** (Just, Woolley & Carpenter, 1982). In a self-paced reading experiment, participants read sentences broken up into segments (words or phrases) by pushing a button that effects the appearance of a segment. In most versions of the task, the previous segments disappear as the presentation of the following segment is initiated, which prevents the rereading of previous segments. Reading times for a segment are analysed as the time between button

presses for the segment to appear and for the following segment. Processing difficulty is analysed as elevated reading time on a given segment in comparison to the same segment in a control condition. Compared to speeded grammaticality judgements, self-paced reading is a more precise measure of on-line reading, although the locality of effects often smears out across several segments (spillover).

The most sensitive behavioural reading-time measure to date is **eye-tracking**. In eye-tracking studies, a high-resolution camera measures movements of the pupil as individuals read texts. Reading times are analysed as the duration of the scans of a particular segment on a line of text. These times are broken up into first pass reading times and second pass reading times, i.e. rescanning of the segment. On top of reading times, ‘first fixation’, i.e. the first time the eyes look at a particular segment, and ‘gaze duration’, i.e. all fixations in a region, are measured as additional indicators of processing difficulty. Processing difficulty is thus measured by reading time as well as number and probability of regressions to a given segment (for overview, see, e.g., Frenck-Mestre, 2005).

The **cross-modal priming** paradigm has mostly been used in studies tapping the (re)activation of antecedents at supposed gap positions. Participants listen to sentences presented via headphones and execute a lexical decision task on visually presented words. The words are visually presented to cooccur at predefined points in the auditorily presented sentences. At the position of the gap, participants are visually presented with either a word that is semantically related to the filler, a semantically unrelated word or a non-word (for variations, see Marinis, 2003). It is assumed that reaction times in the lexical decision task to semantically related words is lower than for semantically unrelated words due to the presence of a gap reactivating the displaced filler.

2.2. Maturational constraints in L2 processing: Neuroimaging and ERP evidence

2.2.1. Neuroimaging

The neurophysiological consequences of putative critical period effects have been studied in the context of L2 processing in neuroimaging and ERP studies. A strong version of the Critical Period Hypothesis would hold that the neural system supporting native-language processing becomes unavailable for late L2 learning, since a critical period restricts the degree of neural reorganization required to accommodate the L2 in cortical areas previously specified for the L1. In neuroimaging studies, strong maturational effects are thus expected to surface in non-overlapping activation of brain areas for non-native and native-language processing. Specifically, native-language processing should engage classical language-related brain areas, whereas L2 processing should activate brain areas

not commonly recruited for language processing. As a correlate, the neurophysiological dynamics of L2 processing should differ from L1 processing.

Earlier studies on covert language production (e.g. Kim et al., 1997), word generation (Price, Green & van Studnitz, 1999), and listening (Dehaene et al., 1997; Perani et al., 1996) report that late L2 learners partially activate different cortical areas compared to natives. By now, the early evidence based on a variety of phenomena, tasks and different participant characteristics (proficiency, age of acquisition) has been supplemented in a fast-growing body of research. Abutalebi, Cappa & Perani (2001), Indefrey (2006b) and Stowe & Sabourin (2005) provide overviews and meta-analyses of recent neuroimaging studies that span production and comprehension as well as different linguistic areas. Abutalebi et al. (2001) conclude that L2s acquired after puberty and to lower proficiency are associated with distinct neural substrates compared to native languages. Importantly, however, when proficiency is held constant, the effects of age of acquisition are mitigated or disappear. Moreover, proficiency appears to be correlated with increased identity in neural activation between adult L2 and native speakers (see also Perani et al., 2003) across tasks and domains. Stowe & Sabourin (2005) offer broadly compatible interpretations on the contingency of neural similarities on proficiency level, adding that some studies on semantic and phonological processing also furnish evidence for recruitment of native-like neural language-associated substrates for low proficient L2 speakers (see also Indefrey, 2006b). Yet, like Abutalebi et al. (2001), they argue that qualitative identity of the brain areas recruited co-occurs with quantitative differences in activation level or spread: Typically, processing the L2 invokes additional activation, which suggests that processing the L2 is in some sense less efficient or automatized as native-language processing.

A set of recent studies adds further evidence to the effects of proficiency and age of acquisition on the processing of syntax and semantics. I selectively discuss three studies and their findings. Using fMRI, Wartenburger et al. (2003) study the neural responses to syntactic and semantic violations in three groups of Italian-German L2 speakers: bilingual learners who have acquired both languages from birth, late L2 learners with low proficiency levels in German, and late L2 learners with a high proficiency level in German as measured in independent general proficiency tasks. The participants read sentences in Italian and German for grammaticality judgements in two conditions. In the syntactic condition, sentences were either grammatical or exhibited morphosyntactic agreement violations (e.g. number, gender or case); in the semantic condition, violations expressed semantic anomaly by role reversal (e.g. *Der Hirsch erschießt den Jäger*, 'The deer shoots the hunter').² Comparisons between activations for judgements in the L1 versus the L2 were run for each condition. Wartenburger et al.

² Note that in this and all the following experiments, the violations labeled 'semantic' do not, in fact, constitute semantic violations as such; rather they express violations of pragmatic plausibility and world knowledge. In keeping with the language-processing literature, I will stick to the label 'semantic'.

(2003) report effects of age of acquisition in the syntactic, yet not the semantic condition. For syntactic judgements, both late learner L2 groups show more extensive neural activity for the L2 in language-sensitive areas, i.e. Broca's area and subcortical structures. The bilingual group did not manifest any language-specific differences in activation. Differences in the strength of cerebral activations depending on proficiency effects are reported for both syntactic and semantic judgements, with the effect being more pronounced in the semantic condition. Again, these differences surface as more extensive activation in language-associated neural structures in L2 processing compared to L1 processing.

In two comparable fMRI studies, Russian-German learners of advanced proficiency levels were scanned while listening to (Rüschemeyer, Fiebach, Kempe & Friederici, 2005) or reading (Rüschemeyer, Zysset & Friederici, 2006) syntactic and semantic violations in the L2. Identically in both studies, syntactic violations were constructed by inserting a preposition followed by a participle (2a); semantic violations denoted violations of plausibility (2b).

- (2) a. *Das Eis wurde im gegessen.
The ice-cream was in-the eaten
b. #Der Vulkan wurde gegessen.
The volcano was eaten

Comparisons were made between neural activation patterns in the L2 group and a native control group. In both experiments, native and L2 subjects show comparable activation of linguistic areas, although with greater accentuation of specific parts of the linguistic network for the L2 group. In the semantic condition, no differences in activation loci or strength between native and L2 speakers occurred in the two studies, irrespective of presentation modality. In the syntactic condition, the L2 groups in both experiments, unlike the native speakers, do not show differences in comparable regions (superior temporal cortex) between syntactically well-formed and ungrammatical sentences; rather these regions were activated and drawn upon irrespective of the grammaticality status of the stimuli. Rüschemeyer et al. (2005) interpret this finding as evidence that lower-proficient L2 speakers commit more resources to syntactic parsing of even simple well-formed structures.

In sum, neuroimaging studies on adult L2 acquirers do not attest fundamental differences in the neural substrates involved in processing late-acquired L2s compared to native languages. Evidence indicates that age of acquisition and, in particular, proficiency level modulate the degree of native-like neural activation patterns in language-related areals, both in terms of the loci and the strength of cerebral activity. Moreover, a higher degree of convergence in neural engagement patterns is reported for semantic processing than

for syntactic processing, suggesting that syntactic processing is more effortful or less efficient for lower-proficient L2 speakers than for native speakers.

2.2.2. Electrophysiological studies

ERP studies explore whether L2ers evoke the same temporal neurophysiological responses as native controls in sentence comprehension. To date, studies have focussed on ERP components associated with semantic integration (N400) and the several syntax-related components (ELAN, LAN, P600). I consider the processing of syntax and semantics and the processing of morphosyntactic agreement relations, in turn.

2.2.2.1. Syntax versus semantics

Following the setup of the Johnson & Newport (1989) study on the acquisition of grammar (Chapter 1.3.2.1), Weber-Fox & Neville (1996) investigated the interaction between the native-likeness of ERP components to syntactic and semantic violations and age of acquisition. Sixty-one Chinese learners of English with a minimum length of residence in the US of five years were recruited and grouped according to their age of arrival (1-3; 4-6; 7-10; 11-13; 16+). Participants read sentences presented word by word for grammaticality judgements. Sentences consisted of semantic violations (3a), syntactic violations (3b), violations of specificity constraints on *wh*-movement (3c) and subadjacency violations (3d).

- (3) a. #The scientist criticized Max's event of the theorem.
- b. *The scientist criticized Max's of proof the theorem.
- c. *What did the scientist criticize Max's proof of?
- d. *What was a proof of criticized by the scientist?

For semantic violations, all groups show N400s in line with native speaker responses. However, delayed latencies on the N400 are observed for learners with ages of onsets of 11 and higher. The ERP results for the syntactic (3b) and the specificity condition (3c) show no early left-anterior negativity for any L2 group and a delayed (11-13) or absent (16+) P600 for the later arrivals.³ Weber-Fox & Neville (1996) conclude that these asymmetries reflect maturational effects on the acquisition of syntax, preempting native-like neurophysiological reflexes for postpubescent learners. A closer look at the data, however, reveals that this interpretation is highly questionable. First, Weber-Fox & Neville report (a) the 11-13 and 16+ group actually did show early, yet bilateral, increases in negativity which the authors dismiss as reflecting a LAN since the younger

³ The ERP data on the subadjacency violation were uninterpretable due to varying baseline effects across groups.

groups do not show them, and (b) that the failure to elicit an ELAN in the L2 groups may be due to the small sample size. More recent research on native speakers' processing of syntactic violations, furthermore, documents that ELANs are not consistently obtained (Osterhout, McLaughlin, Kim, Greenwald & Inoue, 2004), such that the implications of not finding an ELAN in L2 speakers in native-non-native comparisons are unclear. Second, Weber-Fox & Neville (1996: 244) state that age of onset and length of exposure are equally predictive of the P600 amplitude across learners in the syntactic condition. This finding points to the confounding interference of drastic differences in length of exposure as well as proficiency level between the groups. Finally, Weber-Fox and Neville take the lower proficiency levels of the later-arriving groups to be a *consequence* of neurophysiological differences, although it seems equally likely that they might be the *cause* of differences in neurophysiological measures, quite independently of age of arrival.

More recent ERP studies by Hahne and colleagues (Hahne, 2001; Hahne & Friederici, 2001; Rossi, Gugler, Friederici & Hahne, 2006) indeed demonstrate that proficiency level modulates neurophysiological responses to syntax. Using the same stimuli as Rüschemeyer et al. (2005, 2006) in (2), 12 low-proficient postpubescent L1 Japanese learners of German (Hahne & Friederici, 2001) and 16 more advanced adult Russian-German speakers (Hahne, 2001) were tested on semantic and syntactic violations. In both studies, the subjects robustly elicit N400 components for semantic violations. In the syntactic condition, the lower-level L1 Japanese learners show neither an early left anterior negativity nor a later P600; the higher-level L1 Russian learners again show no (E)LAN, yet they do show a P600, signaling sensitivity to syntactic violations. However, since the L1s of the subjects differ between studies, proficiency effects are potentially confounded by cross-linguistic differences.

Rossi et al. (2006) investigated proficiency effects on syntactic ERP signatures by testing 69 German-Italian and Italian-German postpubescent learners who were grouped into high- and low-proficient learners on the basis of a ten-sentence translation task. The main experiment comprised syntactic category (4a), syntactic agreement (4b) and combined category and agreement (4c) violations, here illustrated for Italian.

- (4) a. *Il signore nel beve un caffè.
 The man in-the drinks a coffee
 b. *Il signore nel bar bevo un caffè.
 The man in-the bar drink a coffee
 c. *Il signore nel bevo un caffè.
 The man in-the drink a coffee

ERPs were recorded as subjects listened to sentences word by word and subsequently made grammaticality judgements. The low-proficient L2 groups show qualitative

processing differences compared to high-proficient L2 speakers and to native speakers by demonstrating no (E)LAN for agreement violations and a P600 that is delayed and reduced in amplitude. For the syntactic category violation (4a) and the combined violations (4c), the low-proficient groups do evince an ELAN and a P600, albeit delayed and of smaller amplitudes. By contrast, both groups of high-proficient learners show native-like electrophysiological responses in identical time-windows in all conditions. For the syntactic category and combined violations, the high-proficient groups evince an ELAN, a LAN and a P600; for the agreement violation, they show a LAN and a P600. Subsequent comparisons with a control group of native speakers from Rossi et al. (2005) yield no differences according to language group; only the LAN and the P600, yet not the ELAN component, for category and the combined violations showed larger amplitudes than in native speakers. This quantitative difference is taken to suggest that syntactic analysis is somewhat more effortful and less automatized in high-proficient L2 speakers compared to native speakers.

2.2.2.2. Morphosyntax: Agreement violations

In an ERP study on verbal and gender agreement in L2 Dutch, Sabourin (2003) compared L1 English, L1 German and L1 Romance adult learners with a minimum time of residence in The Netherlands of two years (range 2-50 years). For subject-verb agreement violations, all L2 groups show native-like P600s and late frontal negativities (LANs), albeit both components are somewhat delayed and of lesser amplitude in the L2 subjects compared to the natives. For gender agreement, however, only the L1 German group evinced a native-like P600 effect; the L1 Romance group shows a late frontal negativity; and the L1 English group shows no robust effects at all. These differences point to the effects of cross-linguistic differences in gender marking: Dutch employs a common-neuter gender distinction, Romance languages divide nouns between masculine and feminine gender, and English retains no grammatical gender marking. Finally, German instantiates a tripartite gender system (masculine-feminine-neuter) that overlaps with the Dutch gender system to a large extent in that German feminine and masculine grammatical gender collapse to common gender in Dutch. In all, then, Sabourin's findings suggest that L1-TL proximity in the processing of grammatical gender affects L2 processing, while there are no L1 effects for uniformly realized subject-verb agreement.

Yet, since Sabourin (2003) did not employ an independent measure of proficiency and since knowledge of correct gender was not factored into the ERP analyses, the differences between L1 groups cannot be unambiguously linked to L1 effects. In fact, since the L1 German group's proficiency was considerably higher than the proficiency levels of the other groups, in particular the L1 English group, the group effects may be confounded with proficiency differences. Support for this interpretation comes from recent ERP experiments on gender agreement in L2 French by Foucart & Frenck-Mestre

(2006), who find that L1 English speakers of French do evince ERP signatures of gender agreement violations, although the ERP responses to agreement violations are not fully native-like (see also Osterhout, McLaughlin, Pitkänen, Frenck-Mestre & Molinaro, 2006).

Hahne, Mueller & Clahsen (2006) conducted a study with L1 Russian advanced learners of German on morphological errors in regular and irregular participles and plurals. Research on native speakers (see Hahne et al., 2006 for references) attests differential ERP responses to incorrect regular inflection (LAN and P600) and overregularized irregulars (N400). The L2 subjects are found to evince identical ERP signatures as the natives when reading inflected participles and plurals embedded in sentences. With respect to the morphosyntactically regular inflection, the non-natives showed a LAN and a P600 for incorrect participles, yet only a P600 for incorrect plurals. Hahne et al. (2006) relate the absence of the LAN for plurals to the relatively more complex morphological paradigm of plural formation in German. Finally, Omija, Nakata & Kakigi (2005) report that advanced L1 Japanese learners of English evince a native-like LAN in response to subject-verb agreement errors in English, whereas a lower-proficient group of Japanese learners did not show any effects (see also Chen, Shu, Liu, Zhao & Li, 2007).

2.2.2.3. Summary

Electrophysiological studies that contrast the processing of syntactic and semantic violations resonate with the findings of neuroimaging research in reporting that even lower-proficient L2ers show a robust reliance on semantics, whereas they are less sensitive to syntactic violations, possibly due to less efficient processing. Recent research on highly proficient L2 speakers, however, attests that the full set of native-like processing signatures is replicated in non-native syntactic processing, including early negative components that had previously been frequently reported as missing in lower-level L2 processing (e.g. Hahne, 2001; Mueller, 2006). For highly proficient L2 speakers, these findings suggest that there are no qualitative ERP differences in processing syntactic violations between native and non-native processing. Comparable findings obtain for the L2 processing of morphosyntactic agreement, although the L1 (Sabourin, 2003) and the complexity of the expressions of agreement morphology (Hahne et al., 2006) appear to modulate convergence on TL patterns of brain signatures.

What do these results mean with respect to critical period effects? The more recent evidence from neuroimaging and ERP studies is clearly incompatible with a strong version of the Critical Period Hypothesis that late L2 learners recruit different neural regions and employ distinct neurophysiological mechanisms in language processing compared to native speakers. First, *pace* Weber-Fox & Neville (1996), there is no

evidence of discontinuity in the recruitment of neural regions or processes subserving language that conforms to the geometry of critical periods. Second, late high-proficient L2 speakers show native-like highly automatic brain responses in the L2 processing of syntax and semantics and activation of language areals equally recruited by natives. In conjunction, these findings indicate that classical cortical language areas accommodate late-learned L2s in representation and functional processing, at least at high levels of L2 proficiency. Similar to the behavioural studies on the shape of the age function in L2 acquisition (Chapter 1.3.2.1), then, age of acquisition effects in neurophysiological studies are gradual, starting in early childhood and continue beyond the end of puberty (Abutalebi et al., 2001; Pallier, 2007).

At the same time, L2 processing turns out to be far from identical to native processing. First, syntax is associated with greater non-convergence than semantics, and the degree of native-like cerebral processing of syntax shows greater dependence on proficiency level. Second, L1 effects appear to be implicated at least in the L2 processing of inflectional morphology. These effects suggest that the L1 interferes in L2 processing, thus potentially leading to dissimilarity between non-native and native processing for reasons other than maturation. Finally, quantitative differences can be observed between non-native and native processing. Neuroimaging studies find extended and stronger activation in L2 processing compared to native processing. ERP studies document quantitative differences in amplitude and latency of ERP components even among higher-proficient L2 speakers (Hahne, 2001; Hahne & Friederici, 2001; Rossi et al., 2005; Weber-Fox & Neville, 1996). Such quantitative differences have been argued to reflect lower degrees of efficiency or automaticity in the integration of syntactic category information (Hahne, 2001) or in the reanalysis of the parse (Rossi et al., 2006).

For a better understanding of L2 processing, it is essential to look at aspects like L1 effects and quantitative differences. In L2 processing research, these factors have been predominantly examined in studies employing behavioural measures like reading and reaction times. In contrast to ERP and neuroimaging methods that concentrate on the question whether grammatical information is recruited rapidly and automatically in processing, reading and reaction-time studies also consider whether L2ers recruit the same parsing strategies as native speakers.

2.3. Processing strategies in the L2: Syntactic ambiguity resolution and syntactic dependencies

Most L2 research has been devoted to the cross-linguistic study of the processing of syntactic ambiguities in (a) relative clause attachment, (b) syntactic dependencies in filler-gap constructions and (c) thematic subcategorization. I review these studies to answer the question as to whether late L2ers at advanced proficiency levels converge on native parsing strategies.

2.3.1. Relative clause attachment

Relative clauses like *who never drank wine* in *He thanked the secretary of the professor who never drank wine* preferentially attach to the second NP in English (i.e. low attachment, e.g. Carreiras & Clifton, 1999) but to the first NP in, e.g., Greek, Spanish and French (i.e. high attachment, Cuetos & Mitchell, 1988; Frenck-Mestre & Pynte, 1997). The cross-linguistic differences have been explained as the result of ambiguity resolution strategies in parsing. Various types of models have been advanced that make use of frequency characteristics of the input (Cuetos & Mitchell, 1988), implicit prosody (Fodor, 1998), parametrized syntactic and locality principles in processing (Gibson, Pearlmutter, Canseco-Gonzalez & Hickok, 1996) or a combination of grammatical and pragmatic principles (Frazier & Clifton, 1996) in explaining attachment preferences.

Given that the parsing heuristics underlying the resolution of relative clause attachment ambiguities in native processing are still under debate, L2 processing research has mainly considered (a) whether L2 learners show attachment preferences indicative of the operation of TL ambiguity resolution mechanisms in on-line reading and (b) whether L1 attachment preferences guide relative clause attachment resolution in the TL.

Some studies report that intermediate to advanced L2 learners acquire TL attachment preferences. For Spanish learners of English, Dussias (2003) finds target-like low attachment preferences in the L2, even though these learners, unlike Spanish monolinguals, also show a low attachment preference in their native Spanish (see also Fernandez, 2002; for English-French advanced L2ers, see Dekydtspotter, Edmonds, Liljestrand Fultz, Petrush & Renaud, 2007)). In eye-tracking studies, Frenck-Mestre (2002) finds that highly proficient English speakers have a target-like high attachment preference in L2 French, while a less proficient group shows a low attachment preference as in the L1. For Korean intermediate to advanced learners of Japanese, Miyao & Omaki (2006) report a low attachment preference in reading the L2, which contrasts with the high attachment preference in Korean.

Other studies, however, fail to detect any attachment preferences in L2 learners. Testing advanced learners of English and Greek of various L1s, Felser, Roberts, Marinis & Gross (2003) and Papadopoulou & Clahsen (2003) investigated attachment preferences

both for complex genitives (*the secretary of the professor*), for which attachment preferences are cross-linguistically variable, and for thematic prepositions (*the secretary with the professor*) which universally induce a low attachment preference. For the latter, all non-natives exhibited a low attachment preference in line with native reading patterns. For genitives, though, advanced L2 learners do not show reading-time differences suggestive of any preference, irrespective of whether L1 and TL preferences align (Papadopoulou & Clahsen, 2003) or differ (Felser et al., 2003).

Summarizing, the findings on L2 attachment preferences on complex genitives are inconclusive, with some studies suggesting that L2 speakers exhibit attachment preferences in the L2 that are not transferred from the L1 and others suggesting that L2ers show no sensitivity to either L1 or TL attachment principles. At this point, the reasons as to why some studies fail to record attachment preferences are unclear. The indeterminacy in preferences attested in some studies might arise as a consequence of a transitional stage between L1 and TL preferences for complex genitives in L2 processing (Miyao & Omaki, 2006). L2 learners may pass through a stage where L1 and TL preferences coexist or cancel each other out. In addition, the inconclusive L2 findings might in part reflect a hitherto ill-understood interaction of the factors also claimed to be involved in the native resolution of ambiguities, namely, frequency characteristics of the TL input, effects of L1 and TL prosodic processes in attachment ambiguities (see Dekydtspotter, Donaldson, Edmonds, Liljestrang Fultz & Petrush, in press), or assumptions by L2 speakers about the pragmatic uses of the complex genitive in the TL.

2.3.2. Syntactic dependencies

Studies on syntactic dependencies probe how readers relate displaced constituents, for instance, *wh*-phrases, to their underlying gap position. Reading-time studies on gap-filling effects in English *wh*-movement by Juffs (1998; 2005) and Williams Möbius & Kim (2001) as well as on clitic structures by Hoover and Dwivedi (1998) report that non-natives of various L1s, like natives, attempt to integrate a displaced constituent as early as possible with a potential subcategorizing verb. Marinis, Roberts, Felser & Clahsen (2005) note that these studies do not provide unambiguous evidence of the use of structure-based gap-filling strategies in L2 processing since the design of these studies does not allow for a differentiation between verb-based and structure-based gap-filling strategies.

To distinguish between these accounts, Marinis et al. (2005) look at the reactivation of fillers at intermediate trace positions in *wh*-chains for various L1 groups of advanced L2 English learners. Following Gibson & Warren (2004), Marinis et al. employed self-paced reading to investigate whether the presence of an intermediate trace in (5a) facilitates filler-gap integration at the original trace position.

- (5) a. The nurse who the doctor argued t that the rude patient had angered t is refusing to work late.
 b. The nurse who the doctor's argument about the rude patient had angered t is refusing to work late.

Native speakers of English show filler-gap integration effects at the position following the verb *angered* in both (5a) and (5b); however, the integration effect is significantly reduced for (5a) compared to (5b). It is argued that this interaction of extraction and phrase type reflects the facilitatory effects of intermediate trace reactivation on gap-filling available in (5a), yet not in (5b). Advanced L1 Greek, German, Japanese and Chinese adult learners of English show general filler-gap integration effects at the postverbal position, yet no interaction of extraction and phrase type (for a reanalysis of the data, see Dekydtspotter, Schwartz & Sprouse, 2006). Marinis et al. interpret this finding as evidence that L2 readers do not postulate intermediate traces in processing as a consequence of failing to compute structure-based syntactic representation of filler-gap dependencies.

In a cross-modal priming study following a similar logic, Felser & Roberts (2007) report that advanced L1 Greek learners of English show no reactivation of a filler at the structurally defined original gap position; rather, these learners are argued to integrate the filler with its lexical subcategorizer directly. However, Love, Maas & Swinney (2003) conducted a cross-modal priming task on object-relative clauses such as in (6) that does not support these conclusions. Participants listened to sentences and performed a lexical-decision task on visually presented probes that were either strongly semantically related to the filler item (e.g. *pencil* for (6)) or unrelated (e.g. *jacket*). Superscripted numbers in (6) indicate the probe positions in the sentence.

- (6) The professor insisted that the new exam be completed in ink, so Jimmy used the new pen ⁽¹⁾ that his mother-in-law recently ⁽²⁾ purchased ⁽³⁾ because the multiple colors allowed for more creativity.

In Love et al. (2003), groups of advanced adult and child L2 learners of English with various L1 backgrounds were tested. Although L2 groups show native-like facilitation in lexical decisions for related probes versus unrelated probes at probe point (1), unlike native controls, neither child L2 nor adult L2 learners showed facilitation at the postverbal probe point (3), which Love et al. indicate to mean that, unlike natives, L2 speakers do not reactivate fillers at the postverbal gap position in L2 processing.

Seen in conjunction, studies on the L2 processing of syntactic dependencies find evidence that L2 readers relate fillers to gap positions in on-line reading, yet, the more fine-grained explorations of gap-filling by Marinis et al. (2005), Felser & Roberts (2007) and Love et al. (2003) neither definitively support verb-driven nor structure-based gap-

identification strategies as underlying the processing of long-distance dependencies. Moreover, intermediate trace reactivation in long *wh*-movement could only be found for native speakers with high memory spans, yet not for native speakers with low working memory (Marinis et al., 2005). As the resolution of filler-gap dependencies is thus mediated by cognitive resource restrictions even for natives, the failure of L2 speakers to show native-like performance could be due to these restrictions, rather than divergent processing strategies.

2.3.3. Syntax-argument structure

A third line of reading-time research investigates whether L2 learners are sensitive to thematic properties of verbs in sentence processing. Using eyetracking, Frenck-Mestre & Pynte (1997) examined the processing of verbs that are optionally transitive in English and obligatorily intransitive in French (e.g. *obey* – *obéir*) to see whether the temporary direct object/sentential complement ambiguity of postverbal NPs that holds in English, yet not for the translation equivalent in French, affects L2 processing. In addition, verbs were tested that were obligatorily intransitive in both languages. Two groups of advanced English-French and French-English L2 speakers were compared. Although L2 speakers show additional processing effort on verbs that differ in argument structure between L1 and TL, further downstream they evince processing patterns in accordance with the TL argument structure, irrespective of whether the L1 verbs bear similar or different argument structures compared to the L2 verbs. These findings indicate that L1 argument structure information is initially accessed, but then discarded for TL thematic properties in L2 sentence processing.

Using ERPs, Kotz (1991, cited in Frenck-Mestre, 2005) investigated subcategorization ambiguities as in (7).

- (7) a. The doctor agreed **to** see the patient **had** left the hospital.
- b. The doctor implored **to** see the patient **had** left the hospital.

Sentence (7a) contains an intransitive verb *agree*, and (7b) contains the transitive verb *implore* that requires either a direct object or a sentential complement. Both (7a) and (7b) necessitate reanalysis of the verbs *agree* and *implore* from main-clause verbs to past participles of reduced relative clauses; yet, reanalysis in (7a) and (7b) occurs at different points. Hence, greater processing difficulty is predicted to surface on the preposition *to* for (7b) than for (7a), since *agree* in (7a) can (still) be analysed as the main verb at this point, whereas the verb *implore* in (7b) must be reanalysed as the past participle of a reduced relative clause. Conversely, greater difficulty for (7a) than for (7b) is expected on the verb *had*, which is incompatible with the previous main-verb analysis of *agree* in (7a), yet follows naturally for the previously reanalysed (7b). Both native and advanced

L1 Spanish learners of English demonstrate significantly larger P600 on *to* in (7b) than in (7a), and, conversely, both groups evince larger P600s on the verb *had* in (7a) than in (7b). These findings suggest that the argument structure of TL verbs guides L2 processing (see also Juffs, 1998).

Finally, Montrul (2004) investigated whether advanced L1 English speakers of Spanish show reaction-time reflexes of argument-structure mappings to syntax for Spanish unergative and unaccusative verbs in a visual probe recognition task. Unaccusative verbs project a VS order, and unergative verbs project the canonical SV order; the inverse orders for both verbs are possible, yet require costly syntactic movement (see Chapter 1.3.6.4). The results show an interaction of order and verb type in the response times for the L2 group, which indicates that the argument structure of TL verbs underlies the computation of syntactic structure in L2 processing.

2.3.4. Summary

Broadly in line with neurophysiological studies, the findings of L2 reading-time studies to date indicate that L2 learners make robust use of lexical-thematic information, though potentially less robust use of syntactic information. Nevertheless, the extent of problems in using syntax for ambiguity resolution in the L2 varies across studies. As seen also in ERP data, some studies also indicate that insufficient use of syntax may be related to cross-linguistic transfer and resource limitations in L2 processing. I turn to these factors of L2 processing in the next section.

2.4. L2 processing: Cross-linguistic influence and computational limitations

Most, if not all, studies on L2 processing explicitly rely on non-native versus native speaker comparisons in interpreting the results. Given that L2 acquisition proceeds against the background of successful L1 acquisition, however, L2 processing is likely to be affected by variables specific to L2 development, such as L1 transfer, proficiency differences, etc. In addition, L2 processing is likely to be constrained by procedural limitations, such as insufficient automatization or reduced working memory capacities. The concluding section of this overview of previous research on L2 processing addresses some of these variables, namely, the role of the L1, the role of proficiency and the role of working memory.

2.4.1. L1 effects

Whereas pervasive effects of L1 transfer are attested in the L2 acquisition of grammar, there is less conclusive evidence for the influence of L1 processing strategies on L2 parsing. Within the framework of the Competition Model (MacWhinney & Bates, 1989),

cross-linguistic studies have tested the morphosyntactic and lexico-semantic cues speakers focus on in interpreting sentences. These studies document that learners initially transfer L1 interpretation strategies onto L2 sentences when, e.g., identifying agenthood in off-line tasks (e.g. MacWhinney, 2005). However, since the tasks used in studies within the framework of the Competition Model do not measure on-line processing and concentrate on post-processing conscious interpretation, their results cannot speak directly to the question of whether L1 processing strategies are transferred in the incremental time-course of processing the TL (e.g. Fernandez, 2003).

L1 effects have however been reported in the processing of L2 morphosyntax. Sabourin's (2003) ERP study on Dutch gender agreement processed by L1 English, L1 German and L1 Romance learners of Dutch reports L1-specific differences in that similarity in gender-marking systems between L1 and TL entails greater convergence on L2 processing of nominal agreement (see also Tokowicz & MacWhinney, 2005). In a reading-time study on 'broken agreement' in sentences such as *The key to the cabinets were rusty*, Jiang (2004) reports that advanced Chinese-English L2ers, unlike natives, evince no slowdowns to subject-verb agreement violations, and he suggests that the lack of morphosyntactic number encoding in Chinese underlies the insensitivity to number marking in the L2 (see also Chen et al., 2007).

For syntax, Frenck-Mestre (2002) shows L1 transfer of relative clause attachment preferences in low-proficient English-French learners which gives way to target-like preferences at higher proficiency levels. By contrast, the cross-linguistic studies by Clahsen and colleagues on attachment preferences and trace reactivation do not report any differences between L1 groups. However, although of different L1s, the participants in each study on relative clause attachment did not have L1s that differed amongst one another in attachment preferences (Felser et al., 2003; Papadopoulou & Clahsen, 2003). As for *wh*-movement, the L2 groups tested in Marinis et al. (2005) uniformly fail to show trace-reactivation effects, irrespective of whether the L1s instantiate overt *wh*-movement. Comparing L1 German, Korean and Chinese intermediate learners of English, Williams et al. (2001) find no difference in garden-path patterns in on-line reading times and plausibility judgements in *wh*-questions. In a partial replication of Williams et al. (2001), Williams (2006) reports analogous null effects of L1 differences on the incremental computation of plausibility for syntactic reanalysis in filler-gap structures for advanced L1 Chinese and L1 Romance speakers of English. In conjunction, these studies suggest that the parsing of *wh*-structures in the L2 is not modulated by the availability of overt *wh*-movement in the L1.

The limited evidence available to date thus suggests that L1 transfer likely differentially affects grammatical domains in L2 processing. Effects of the L1 on L2 processing surface for morphosyntactic agreement and also argument structure (Frenck-Mestre & Pynte, 1997), while the majority of studies on the L2 processing of syntax

cannot detect L1 effects. Clearly, further research is necessary to see whether this pattern can be generalized.

2.4.2. Proficiency, automaticity and working memory in L2 processing

2.4.2.1. Proficiency

On top of L1 effects, L2 processing is likely to be modulated by the degree of acquisition of the TL. Across paradigms, the degree of native-like processing has been found to be correlated with proficiency level. In neuroimaging studies (e.g. Perani et al. 2003) and ERP studies on the processing of morphosyntax (e.g. Ojima et al., 2005; Rossi et al., 2006; Weber-Fox & Neville, 1996) proficiency differences underlie substantive contrasts in the recruitment of native-like brain areas and electrophysiological responses. Moreover, ERP studies on the processing of artificial grammars (e.g. Friederici, Steinhauer & Pfeifer, 2002) and miniature natural-language grammars (e.g. Mueller, 2006; Mueller, Hahne, Fujii & Friederici, 2005) demonstrate that, even after a few hours of training on toy grammars, learners develop processing components, such as the P600 and LANs, in response to syntactic violations that resemble ERP components found in native-language processing. For natural-language acquisition, Osterhout et al. (2006) report analogous developments of syntax-related ERP components in English learners of French after only several months of exposure. Both semantic violations and morphosyntactic violations in subject-verb and determiner-noun agreement were presented to the L2 learners at one month, 4 months and 8 months of university classroom instruction in French. Semantic violations were found to elicit N400s after one month of instruction. At the same time point, L2 learners equally evinced N400s in response to morphosyntactic violations, which documents that they differentiated between well-formed and ill-formed syntax, albeit not in a native-like manner. After 4 months, this latter N400 was supplanted by a native-like P600 indexing syntactic violations. In a similar vein, longitudinal neuroimaging research on native Mandarin Chinese speakers acquiring Dutch (Indefrey, 2006b) provides evidence that syntax-specific brain regions implicated in native-language processing are recruited after only a few months of TL exposure. At the initial stages of L2 processing explored in these studies, increasing proficiency in the L2 implicates greater acquisition of the grammar of the TL which can then be recruited in L2 processing.

Effects of proficiency on processing patterns are also seen in higher-proficient L2 learners that have successfully acquired the domain of the TL grammar under investigation. As discussed in Section 2.2.2.1 above, Hahne (2001) and Hahne & Friederici (2001) administered the same phrase-structure violations involving a preposition followed by a verb (e.g. *The ice cream was in the eaten*; see (2a)) to an intermediate L1 Japanese group which had been exposed to German for a mean of 2.5

years and a more advanced L1 Russian group which had a mean exposure to German of 5 years. Off-line behavioural data confirmed that both groups correctly judged these sentences, which indicates that they have acquired grammatical knowledge of the syntactic violation. However, of the two groups, only the more advanced L1 Russian group evinced a P600 in response to syntactic category violations. The study by Rossi et al. (2006) on similar materials reviewed above further demonstrates that low and high-proficient Italian-German and German-Italian late L2 learners differ in that the high-proficient groups evince ELANs and LANs for syntactic category violations absent in the lower-proficient groups. These studies indicate that proficiency differences at higher levels incur substantive changes in processing, even though L2 grammatical acquisition of the phenomena at hand is not at stake. In these cases, increased proficiency likely reflects more automatized processing of the L2 and more efficient deployment of cognitive resources in L2 processing.

2.4.2.2. Automaticity

A typical feature of non-native processing is that it is slower than native processing. The slower speed of L2 processing could potentially serve as an index of divergent processing. However, processing speed *per se* does not appear to account for processing differences. For instance, Hoover & Dwivedi (1998) contrasted the reading pattern of the French sentences in (8) in a self-paced reading task.

- (8) a. Sarah le fera signer en presence de son avocat.
 Sarah it will-have signed in the presence of her lawyer
 b. Sarah le fera demain en presence de son avocat.
 Sarah it will-do tomorrow in the presence of her lawyer

Due to the ambiguity between *faire* ('make') as a thematic verb (8b) and a causative verb (8a), the clitic in (8a) is initially misanalysed as the argument of the immediately following verb *fera* by native French speakers. This is borne out by the subsequent processing difficulty when the clitic needs to be reanalysed as the argument of the verb *signer* ('sign'). A group of highly proficient L1 English speakers of French was subdivided into slow and fast readers on a separate reading test in order to address the question whether sensitivity to syntactic ambiguities is modulated by speed. Native French speakers were used as controls. The results show no interaction between group (native, fast L2, slow L2) and ambiguity resolution, since all groups showed analogous slowdowns on the verb *signer* compared to the temporal adverb *demain* ('tomorrow'). Even though both L2 groups are significantly slower than natives and the L2 groups differed in reading speed amongst each other, speed *per se* was not a factor predicting differences in syntactic processing (see also Frenck-Mestre, 2002; McDonald, 2006).

On other views, procedural limitations in L2 comprehension have been categorized under the rubric of ‘automaticity’. Automaticity denotes the routinization of cognitive tasks that do not involve attentional control. Crucially, automaticity is not linearly dependent on speed; rather, it indexes the replacement of less efficient, parallel or redundant subcomponents, e.g. in language comprehension, by more efficient and integrative processes (for extensive discussion, see Segalowitz, 2003; Segalowitz & Hulstijn, 2005).

In electrophysiological studies, differences in automaticity between native and L2 speakers have been linked to qualitative differences in the context of early ERP components (ELAN) as well as in quantitative differences in amplitude and latency of ERP components in grammatical processing. As the previous sections illustrated, quantitative differences in amplitude size and latency of ERP components associated with semantic (N400) and syntactic (P600) processes between native speakers and especially lower-level L2 speakers are frequently attested. In neuroimaging studies, differences between adult non-native and native speakers in strength and spread of qualitatively overlapping neural substrates have also been interpreted as reflecting decreased efficiency and increased effort in non-native processing (e.g. Stowe & Sabourin, 2005; Wartenburger et al., 2003).

At the behavioural level, a study by Kilborn (1992), for instance, provides evidence that resource limitations constrain the automaticity of L2 sentence processing even at high proficiency levels. In a word-monitoring task, participants had to identify target words embedded in sentences in three conditions: (1) ‘normal’ sentences that included syntactic and semantic information, (2) ‘syntactic’ sentences in which all open class words were replaced with random words of the same word class, so that syntactic relations were maintained but semantic information was lost, and (3) random strings of words which did not maintain syntactic or semantic relations. By gradually removing different types of information in the same task, the relative contributions of individual information types were explored in on-line reading. Two participant groups, native speakers of English and highly proficient L1 German speakers of English, performed the task in two modes. The native speakers performed the task under normal conditions and under increased task demands, that is by listening to the stimuli partially masked under pink noise. The L2 group performed the task under normal conditions in their L1 (German) and the L2 (English). The mean reaction times for each condition are given in Table 2.1.

Stimulus Type	English L1 speakers		German-English L2 speakers	
	Clean	Noise	German (L1)	English (L2)
Normal sentences	230	322	221	292
Syntactic sentences	304	361	264	317
Random strings	341	393	342	354

Table 2.1. Mean reaction times (in milliseconds) in word monitoring task (adapted from Kilborn, 1992).

Native English speakers and L2 group performing the task in their L1 show significantly faster response times in word monitoring when syntactic information is available compared to randomly presented strings of words and further facilitation when additional semantic information is supplied. Hence, native speakers use syntactic information and they rapidly integrate syntactic and semantic information which leads to faster word monitoring times than when only syntactic information is available.

By contrast, high-proficient German-English speakers show facilitation for syntactic information compared to random sentences, yet no statistically significant further processing benefit for additional semantic information (292ms versus 317ms). The lack of facilitation indicates that additional semantic information fails to be integrated with syntactic information to facilitate word monitoring (yet, see Scherag, Demuth, Rösler, Neville & Röder, 2004). Strikingly, native processing resembles the non-native pattern when natives perform the task under increased task demands (pink noise). In comparisons of the natives' performance between the clean and noisy condition, a group-by-stimulus type interaction indicates that normal sentences, requiring the integration of syntactic and semantic information, elicit the greatest disruption of word monitoring under noise. This interaction suggests that the increased task demands particularly affect the automatic integration of syntactic and semantic information in on-line reading. Taken together, these findings (a) point to limitations in computational resources for integrating multiple types of information for non-natives and (b) suggest that analogous limitations can be forced for natives under stress.

Such limitations have also been shown to affect behaviour in off-line grammaticality judgements. In McDonald (2006), a subset of the items from Johnson & Newport (1989) – discussed in Chapter 1.4.2.1 – involving word order violations in declaratives and questions as well as ungrammatical inflectional morphology were tested on intermediate to advanced postpubescent L2 speakers of English from a variety of L1 backgrounds. In addition to an auditory grammaticality judgement task, the participants completed (a) a gating task to test lexical decoding ability, (b) a working memory task, and (c) a word detection task to test processing speed. The results of the grammaticality judgement task replicate the findings from Johnson & Newport (1989) and its successors in that word order is judged more accurately than inflectional morphology (see Chapter 1). Importantly, for the L2ers, accuracy in judgements is significantly correlated with

lexical decoding ability as measured in the gating task and working memory span. These correlations indicate that L2 performance in off-line judgements is affected by extragrammatical resource limitations. In a second step, native speakers of English performed the same judgement task when subjected to different stressors, i.e. (a) concurrent memorization tasks, (b) listening to stimuli masked by white noise, (c) listening to stimuli at compressed speech, and (d) performing judgements under response time pressure. Correlational analyses show that natives, who perform at ceiling under unstressed conditions, mirror the non-native pattern of differentially decreased judgement accuracies across constructions when listening to stimuli under white noise.

In sum, the findings from Kilborn (1992) and McDonald (2006) document (a) that non-native off-line and on-line performance is highly susceptible to resource limitations and (b) that forcing resource limitations in natives may lead to L2-like performance patterns. In conjunction, they adduce evidence that non-target-like L2 behaviour does not necessarily index lack of linguistic knowledge or grammatical deficits, but may indicate that L2 learners process the TL in an analogous way to native processing under increased task demands. The lower automaticity of L2 processing points to restrictions in the computational resources available in L2 processing. Such restrictions have been conceptualized in terms of different working memory capacities.

2.4.2.3. Working memory

Working memory has been shown to affect both the application of parsing strategies and the outcomes in sentence processing for native speakers. Sentence processing requires the parser to identify, hold in memory and integrate different kinds of information, e.g. morphological, syntactic, semantic and discourse information. This complex cognitive task is constrained by the capacities of short-term memory (for general overview, see Baddeley, 2006; Caplan & Waters, 1999). For the processing of complex sentences, King & Just (1991) show that native speakers with a low memory span take longer to read the complex regions than high-span speakers. For syntactically ambiguous sentences, e.g. Just & Carpenter (1992) and MacDonald, Just & Carpenter (1992) find that low-span natives show qualitatively different processing strategies compared to high-span natives. High-span readers appear to compute multiple analyses of ambiguous structures and keep them in memory, while low-span readers quickly commit to the simpler analysis (for critique, see, e.g., Caplan & Waters, 1999).

In ERP studies on syntactic ambiguity resolution, qualitative and quantitative differences between high-span and low-span native speakers have been reported by, e.g., Friederici, Steinhauer, Mecklinger & Meyer (1998) and Bornkessel, Fiebach & Friederici (2004). In these studies, high-span natives demonstrate P600 effects for syntactic reanalysis in ambiguous sentences, but low-span participants show N400 effects. This

difference is taken to suggest that the low-span participants do not effect syntactically-based reanalysis (see also Vos, Gunter, Schriefers & Friederici, 2001). For natives, interindividual processing differences are thus partially related to restrictions of processing resources.

In an L2 context, interindividual differences in working memory have been found to affect global L2 proficiency measures (Harrington & Sawyer, 1992; Miyake & Friedman, 1998), lexical retrieval and vocabulary use (e.g. Kroll, 2006) as well as text comprehension (Walter, 2004). With respect to syntactic processing, however, no robust effects of working memory on parsing strategies have thus far been attested.

In a replication of Juffs & Harrington's (1995) study on subject- versus object-extraction in *wh*-movement, Juffs (2004; 2005) tests for correlations between working memory scores on reading span and word span memory tasks and the (re-)analysis of garden-path sentences for L1 Chinese, Japanese and Spanish intermediate learners of English. Specifically, it was tested (a) whether working memory scores predict differences for reading times on the critical region of disambiguation and (b) whether working memory interacts with sentence type, i.e. garden-path and non-garden-path sentences. Although low-span L2ers read more slowly than the high-span L2ers overall, there is neither an effect of working memory on the critical segment nor an interaction of working memory and sentence type.

Replicating earlier findings (Swets, Desmet, Hambrick & Ferreira, 2007), Omaki (2005) reports that relative clause attachment preferences vary for English natives according to memory span. However, in subsequent experiments that use the same materials for L1 advanced Japanese learners of English, working memory in the L2 as measured in a reading-span task does not correlate with relative clause attachment preferences either off-line or on-line. There is tentative evidence in Omaki (2005), though, that the L2 speaker's processing preferences resemble those of low-span natives. Other studies, e.g. a cross-modal priming study on trace activation in filler-gap dependencies (Felser & Roberts, 2007), find that high-span L2ers show processing patterns distinct from both high-span and low-span natives.

In all studies investigating working memory differences among L2 speakers and between native and non-native speakers, L2 speakers generally achieve lower memory scores than natives, at least in verbal working memory tasks. Moreover, the working memory spans for high-span L2 subjects overlap with low-span natives, and not with high-span natives (Felser & Roberts, 2007; Juffs, 2005; Omaki, 2005; Williams, 2006).⁴ In a direct comparison of non-native and native speakers on working memory in the TL measured in a syntax-independent recall task, McDonald (2006) reports highly significant differences between L2 speakers and natives. It thus seems that working memory

⁴ The L2 subjects in Felser & Roberts (2007) are compared with native subjects in Roberts, Marinis, Felser & Clahsen (2007) who took a different reading span test, such that cross-study comparisons are only possible to a limited extent.

resources are strained by processing the L2 *per se*, rather than solely by interindividual memory-span differences within L2 processing (Indefrey, 2006a). Indeed, a study by Service, Simola, Metsänheimo & Maury (2002) compared working memory resources in sentence processing for highly proficient L1 Finnish speakers of English in the L1 and the L2. Participants judged whether declarative sentences matched pictures that were simultaneously shown to them; at the same time, they memorized the final word of the sentences. Sentences were presented in sets growing in number. Working memory was measured as the largest number of sentences with correct judgements and recall. The participants show consistently smaller working memory scores, lower judgement accuracy and longer judgement latencies for English sentences than for Finnish sentences. Even when one of the dual tasks, i.e. judgement or recall, is removed, performance in the L2 falls significantly below the levels of performance in the L1. Further, L2 working memory span positively correlates with accuracy and (negatively) with reaction times of sentence judgements. Service et al. (2002) argue that the lower automaticity of L2 sentence processing compared to L1 processing places larger demands on working memory resources and thus constrains L2 sentence processing. In this perspective, research on working memory ties in with the findings on the automaticity of L2 processing reported in the sections above that construe resource limitations as a general trait of L2 processing.

In sum, the individual contributions and interactions of factors affecting processing in non-native languages, be they the influence of the first language, L2 proficiency, integration difficulty or memory constraints, still remain to be ascertained. The limited amount of research available to date finds that all of these factors affect at least certain domains in L2 processing. It is conceivable that (some of) these factors conspire to evince surface deviance in L2 processing that bears the hallmark of substantial dissimilarity between non-native and native processing. To date, the scope and interactions of these factors are far from well understood.

2.5. Methodological limitations

Since the study of L2 processing is a comparatively new line of research, many studies invite methodological criticisms in terms of their design, their participant selection and the validity and generalizability of their findings. Rather than offering individual critiques of the studies reviewed above, I limit myself to noting some general conceptual and methodological problems common to a number of studies.

First, much research on L2 processing is couched within the terms of the critical period debate that seeks to address the question whether adult L2ers converge on native-language processing sources and resources. However, not a single study has thus far

considered near-native L2 speakers at ultimate attainment to test the limits of L2 processing. Given that proficiency level in the L2 has been shown to correlate with substantive changes in processing across methods and tasks, it is a somewhat surprising gap in research that no systematic comparisons have been undertaken between endstate L1 and endstate L2 speakers. The logic of comparisons between non-native and native speakers dictates that finding lower-proficient L2 learners who show particular native-like processing reflexes indeed allows for the generalization of these findings to higher-proficient L2 learners, whereas the finding that lower-proficient, intermediate or advanced L2 learners do not evince certain native processing effects simply does not allow for any inferences about the upper limits of L2 processing.

Second, research on L2 processing to date has given short shrift to cross-linguistic differences. Although the number of studies that compare and contrast L2 speakers with different L1 backgrounds is increasing, very little research in L2 processing considers properties that are systematically different between the L1s as well as between the L1s and the TL. Existing studies compare groups of L2 speakers whose L1s are identical for the phenomenon under investigation (e.g. Felser & Roberts, 2007; Papadopoulou & Clahsen, 2003). Other studies that do include L1 groups that differ among one another in relation to TL properties only allow for limited conclusions about the extent of L1 effects, because the L1 groups are not sufficiently matched on proficiency (e.g. Sabourin, 2003).

Third, generalizations about L2 processing are usually made across several studies. So far, there have been very few studies, if any, testing a range of interconnected phenomena across several linguistic domains to arrive at a more comprehensive understanding of L2 processing for a single or comparable set of L2 subjects. On the basis of cross-study comparisons, it is uncertain whether differences reported for L2 processing are ultimately due to meaningful differences in the architecture of L2 processing or to differences in, e.g., participant selection between studies.

2.6. Summary of L2 processing

The evidence emerging from the growing body of research on L2 on-line sentence comprehension cautions against sweeping generalizations about L2 processing and its (dis-) similarity compared to native-language processing. Whereas there is robust evidence across experimental paradigms that L2 speakers converge on the TL in the semantic and lexical-thematic domains, a more complex picture emerges for the processing of morphosyntax. Several studies report differences between non-natives and natives in the processing of morphosyntactic agreement violations; however, they also supply evidence that dissimilarity is modulated by L1, L2 proficiency and task demands. For syntax, ERP studies demonstrate convergence on native-speaker signatures in the

processing of syntactic violations, at least for highly proficient L2 speakers. Reading-time research, however, finds differences between advanced non-native and native processing of, e.g., long-distance dependencies and attachment ambiguities. At the same time, some of the few studies in this domain also point to modulation of the degree of differences by proficiency level and constraints on working memory.

With a view to the issue of asymmetries between non-native and native processing, let us return to the potential differences between L2 and L1 processing listed by Clahsen & Felser (2006b) and repeated in (9) to assess them briefly against the evidence available to date.

- (9=1) a. L2 processing is fundamentally different from native processing
L2 acquirers recruit different processing mechanisms and subservient neural regions compared to native speakers.
- b. L2 processing only makes partial use of native processing mechanisms
L2 acquirers do not recruit the full set of processing mechanisms or neural substrates compared to native speakers.
- c. L2 processing is affected by L1 transfer
The interference of L1 processing mechanisms entails differences between non-native and native processing.
- d. L2 processing is computationally less efficient than native processing
L2 acquirers cannot process the TL as quickly and as automatically as native speakers.

As discussed above, recent research, in particular on high-proficient L2 speakers, yields no evidence to support the position in (9a) that there are fundamental differences between non-native and native processing in terms of either the processing mechanisms tested in ERP and reaction-time research or the neural substrates of L2 processing probed in neuroimaging studies. However, differences between non-native and native processing patterns in reading-time studies on the resolution of attachment ambiguities and long-distance dependencies have been interpreted as evidence in favour of critical period effects affecting specific domains of language processing (9b). I review a proposal, the ‘Shallow Structure’ Hypothesis (Clahsen & Felser, 2006a; 2006c) claiming that the use of abstract grammar is restricted in L2 processing by a critical period (Section 2.7). Another model to be reviewed, the ‘Declarative-Procedural’ Model (Ullman, 2005) argues that late L2ers make use of different mental processing pathways compared to natives as a consequence of a neurophysiological critical period. L1 transfer (9c) has been shown to affect L2 processing of morphosyntactic agreement relations, yet there is little evidence of L1 transfer in syntactic processing. As for (9d), there is converging evidence that L2 processing is slower and less automatized than native-language processing. However, it is

an open question whether limitations in automatization mean that the same processes are implicated in non-native and native processing, although they are carried out less efficiently in the L2, or whether resource constraints lead to more taxing computations not being fully executed or not being executed at all in the L2. One proposal that appeals to computational differences between native and non-native processing, the ‘Capacity’ model (McDonald, 2006), is reviewed in Section 2.7.3.

2.7. Models of adult L2 processing

2.7.1. Grammatical impairment: The ‘Shallow Structure’ Hypothesis

Clahsen & Felser's (2006a; 2006b; 2006c) ‘Shallow Structure’ Hypothesis takes as its starting point the symmetries between non-native and native processing of morphology, (lexical) semantics and thematic structure as opposed to asymmetries in the processing of syntax. Based on the findings that intermediate to advanced L2 speakers of various L1s (a) fail to evince syntactically-based attachment preferences in the on-line comprehension of ambiguous relative clauses (Felser et al., 2003; Papadopoulou & Clahsen, 2003) and (b) fail to show trace (re-) activation effects in the processing of syntactic dependencies (Felser & Roberts, 2007; Marinis et al., 2005), Clahsen & Felser (2006a; 2006b; 2006c) argue that adult L2 learners employ different ‘shallow’ processing strategies lacking syntactic detail compared to natives. L2 parsing is argued to be shallow by capitalizing on lexical-thematic and semantico-pragmatic information at the expense of computing abstract phrase structure.

According to Clahsen & Felser (2006a; 2006b; 2006c), the use of shallow processing is a matter of degree in that L2 speakers are said to be ‘largely restricted’ to the shallow processing route, a parsing route in principle available to native speakers, too (e.g. Townsend & Bever, 2001, see below). For L2 speakers, however, the use of shallow processing is deemed to be forced by divergent grammatical representations that constrain L2 acquisition as a result of a critical period in syntax.

[T]here is evidence that second-language learners develop interlanguage grammars that are fundamentally different from L1 grammars (e.g. Bley-Vroman, 1990; Clahsen & Muysken, 1986, 1989). In short, we believe that although both processing routes are available to L2 learners in principle, successful structural parsing depends on the availability (and accessibility) of sufficiently detailed implicit grammatical knowledge. [...] With the full parsing route being of limited use in L2 processing, learners’ interpretations will typically be derived via the shallow processing route only. [...] Under this view, whether or not L2 learners can also develop native-like parsing abilities will depend on their acquiring a native-like grammar. (Clahsen & Felser, 2006a: 117-118)

On the assumptions that non-convergence in L2 grammars forces shallow processing, L2 speakers are argued to analyse complex nominals such as *the secretary of the professor* as a chunk lacking internal structure and to process filler-gap structures by directly integrating the filler with its lexical subcategorizer, thus failing to evince native-like attachment preferences and trace reactivation. The ‘Shallow Structure’ Hypothesis claims that L1 transfer effects in the realm of syntactic processing do not surface since transfer of L1 syntactic processing strategies would presuppose the availability of grammatical architecture in the L2.

The ‘Shallow Structure’ Hypothesis is based on a wide-ranging synthesis of experimental findings in L1 and L2 processing research. Grounded in psycholinguistic theory, it offers a unitary explanation for dissimilarities between syntax and other linguistic components in processing and provides a clear perspective on effects of L1 transfer and cognitive resource limitations in L2 processing.

However, the explanatory scope of the hypothesis is constrained by the vagueness of its central claim that L2 speakers are largely restricted to shallow processing but can use phrase-structure based parsing in principle. This leaves open the question as to whether quantitative differences between non-native and native processing sets non-native processing apart from native processing or constitutes a developmental stage that – provided proficiency, L2 exposure and usage, working memory capacities etc. increase – gives way to native-like processing. Ultimately, this is an empirical question, as conceded by Clahsen & Felser (2006a), yet, these open issues compromise the predictive scope of the ‘Shallow Structure’ Hypothesis. Moreover, if shallow processing is an option available to native speakers as well, who, by definition, possess full-fledged implicit grammatical representations that are suitable for parsing, it is not clear why and when a shallow parsing route in non-natives reflects divergent underlying grammatical representations and when it resembles native-like recourse to a shallow route in lieu of an also available syntactic parsing route.

Conceptually, it seems problematic to locate the causes of processing deviance in divergent grammatical representations even though non-natives show target-like off-line behaviour in the experiments carried out by the Clahsen and Felser group. Such an explanation can only be upheld by dismissing the suitability of off-line tasks for determining whether “the nature and extent of their grammatical knowledge was native-like” (Clahsen & Felser, 2006a: 120) and privileging on-line processing data. However, if there is no way of establishing grammatical knowledge other than inferring it from on-line data, the ‘Shallow Structure’ Hypothesis runs the risk of becoming circular and immune to falsification.

Empirically, the actual data set on which the theory is based is fairly small and comprises solely two phenomena, i.e. relative clause attachment and *wh*-movement, for which conflicting results have been obtained (see Sections 2.3.1 and 2.3.2 above). In addition, the Shallow Structure explanation for the data on which the hypothesis is based

does not seem convincing. For instance, the contention seems questionable that the failure of L2 speakers to employ intermediate traces in long-distance *wh*-extraction is due to a non-native-like grammatical representations. Long-distance *wh*-extraction of the type investigated by Marinis et al. (2005) is constrained by Subjacency, a grammatical principle tied to the representation of intermediate traces (Chomsky, 1981). According to the ‘Shallow Structure’ Hypothesis, L2 learners’ grammatical representations on Subjacency should be deviant, causing deviant processing. However, a large body of research has shown that advanced L2 learners of various L1 backgrounds show robust knowledge of Subjacency constraints on *wh*-movement (for overview, see, e.g., Hawkins, 2001a). Moreover, since these constraints are underdetermined in the input, they cannot be learnt explicitly or by analogy to the L1. This leads to the conclusion that abstract grammatical knowledge is implicated in L2 grammatical representations of long-distance *wh*-movement. It seems an ad hoc stipulation that such grammatical knowledge cannot be used for parsing in some sense.

In light of these points, much more research on on-line parsing of syntax and related off-line knowledge is necessary to substantiate the claims of the ‘Shallow Structure’ Hypothesis that grammatical divergence incurs processing differences between non-natives and natives. Seeing that the hypothesis imposes constraints on the potential of L2 processing, it seems apt to investigate the limits of L2 processing capabilities with L2 endstate speakers to factor out effects of proficiency for L1-L2 processing comparisons.

2.7.2. Different mental processes: Procedural versus declarative knowledge

Other approaches construe differences between monolinguals and L2 speakers in terms of the mental routes for accessing and retrieving grammatical knowledge (e.g. DeKeyser, 2000; Ellis, 2002; Jiang, 2004; Paradis, 1997; 2004; Ullman, 2005). Unlike models assuming differences in grammatical representation between L1 and L2 acquisition (see above and Chapter 1.4.3.1), these models focus on how linguistic knowledge is applied in language comprehension and production. Common to these models is a broadly similar distinction between two types of linguistic knowledge accessed via two associated processing routes, one being subject to attentional control and conscious inspection, and the other being applied automatically and unconsciously. This distinction is variously referred to as explicit versus implicit knowledge (DeKeyser, 2000; 2003; Hulstijn, 2002; Paradis, 2004), non-integrated versus integrated knowledge (Jiang, 2004; 2007), or declarative versus procedural knowledge (Ullman, 2005). According to these approaches, L1 processing of syntax is largely automatic, drawing on implicit, integrated knowledge. In contrast, L2 processing is restricted to the use of explicit, non-integrated knowledge, although its application may be speeded up to degrees resulting in high levels of L2 proficiency and fluency.

One of these models, Ullman's 'Declarative/Procedural' (DP) model, builds on two distinct neural and mental pathways of representing and processing language: the declarative system and the procedural system (Ullman, 2001; 2004; 2005).⁵ The declarative system is in charge of learning and processing facts, events, words, etc., and the procedural system is responsible for the acquisition and control of automatic cognitive and motor skills. Each of the systems is subserved by a specific neurophysiological circuitry (for details, see Hartshorne & Ullman, 2006; Ullman, 2001; 2005). For linguistic processing, the declarative system is equated with lexical storage and encompasses explicit and (some) implicit knowledge; the procedural system underlies rule-governed (hierarchical) computations such as syntax or regular morphology and draws on implicit knowledge only. In native language processing, both systems interact and overlap. For adult L2 acquisition, Ullman contends that neurophysiological changes as the consequence of a critical period lead to the attenuation of the procedural system whilst leaving the declarative system largely unaffected.

Since both the declarative and the procedural system are available in native processing, maturation does not lead to fundamental differences in processing mechanisms or neurophysiological architecture between native and L2 processing; rather, it forces differences in the extent to which the two mental processing systems are recruited. According to Ullman, adult L2 acquirers predominantly rely on declarative knowledge, storing chunks and formula to compensate for the inefficiency of the procedural system. Hence, lower-proficient learners are supposed to exhibit particular difficulties with complex syntax, e.g. long-distance dependencies that do not lend themselves to memorization. After prolonged exposure to the TL, the procedural system in L2 speakers is supposed to develop, take over functions thus far executed by the declarative system and lead to improved performance on, e.g., complex syntax and non-local dependencies. In fact, depending on the nature and duration of TL exposure and intraleaner factors such as sex, age, learning potential, etc., acquisition of the procedural system underlying grammatical processing is argued to be possible to native-like levels (see also Green, 2003).

The DP model's primary strength lies in its strong neurophysiological orientation and grounding in cognitive psychology. With respect to ultimate attainment, the DP model suggests that native-likeness can emerge as a function of prolonged exposure and high proficiency in the L2 that leads to augmented efficiency and automaticity in the procedural system. As noted by Ullman (2005) in this respect, the DP model runs counter to the categorical predictions of the Critical Period Hypothesis because it does not posit a qualitative difference enforced by maturational changes. Rather, the model emphasizes the relevance of exposure and proficiency as determinants for convergence on native

⁵ I limit the discussion in this section to the DP model as an approach exemplifying the logic of similar and related models, such as Paradis' neurolinguistic theory of bilingualism (Paradis, 2004).

processing in terms of the automaticity of the underlying processes and the involvement of neurophysiological components.

However, the wide-ranging scope of the model does not seem to allow for precise predictions regarding L2 processing. First, the gradient notion of ‘reduced availability’ leaves open whether the procedural system is partially defunct or non-existing in adult L2 acquisition or whether it is fully instantiated, yet operates at reduced efficiency (see also Clahsen & Felser, 2006b). Second, the model’s broad scope and lack of linguistic detail leave open how the DP distinction relates to grammatical theory, and, indeed, whether grammatical theory has any independent status in the DP model. For L2 acquisition, this lack of detail makes it difficult to assess whether a given grammatical phenomenon is exclusively effected in the procedural system and should thus be hard or impossible to master in the declarative system for L2 speakers. Third, the composite of factors that are argued to engender the transition from the declarative to the procedural system makes it hard to pin down the prerequisites for the emergence of native-like processing mechanisms in adult L2 speakers. In this regard, the DP model also raises questions about continuity in L2 processing. Moreover, it is unclear whether factors other than input frequency, such as working memory or L1 transfer, can affect L2 processing and, if so, how they interact with the two systems. It will thus be of interest to study whether L1 differences modulate ultimate attainment in L2 processing.

2.7.3. Capacity models of sentence processing

Finally, a third type of approach holds that differences between non-native and native language processing do not reflect representational deficits in grammar or different processing architectures, but are due to limitations in computational resources in L2 processing. Unlike the two previous approaches, capacity models do not postulate any critical period effects for L2 processing.

A way of testing whether non-target-like performance is related to computational capacity limitations is to ascertain if such performance can be reproduced in normal monolingual adults by subjecting them to exogenous processing burdens that simulate the endogenous processing constraints presumed to characterize, e.g., L2 processing. Exogenous processing conditions are manipulated by masking the input signal (e.g. by noise, low-band filtering, etc.), by speeding up the processing task or by adding concurrent tasks to sentence processing in order to engender cognitive overload. Studies adhering to this rationale by Kilborn (1992) and McDonald (2006) find that processing and judgement patterns typical of L2 performance surface for monolinguals performing under stress (Section 2.4.2.2). Further, native adults have been compared to children with specific language impairment (e.g. Hayiou-Thomas, Bishop & Plunkett, 2004) and, in particular, agrammatic aphasics (e.g. Caplan, Waters, DeDe, Michaud & Reddy, 2007; Dick et al., 2001; Miyake, Carpenter & Just, 1994). For morphosyntactic agreement,

normals performing under noisy (Kilborn 1991) or dual-task conditions (Blackwell & Bates, 1995) were found to evince decreased reliance of agreement markers in sentence interpretation in a manner similar to aphasics. For complex syntax, Miyake, Carpenter & Just (1994) and Dick, Gernsbacher & Robertson (2002) equally elicit selective difficulties in comprehension under speeded presentation or dual-task load that mimic the problems with non-canonical word orders typical of Broca's aphasics.

In light of these findings, capacity models face the question as to what type of processing strategies apply for L2ers and natives when resources are taxed or overburdened. A possible characterization of processing strategies under load is given by two-stage models of (native) sentence comprehension, such as the 'Late Assignment of Syntax Theory' (LAST) by Townsend & Bever (2001) and the approach of 'good enough' processing by Ferreira and colleagues (e.g. Ferreira, 2003; Ferreira, Ferraro & Bailey, 2002; Ferreira & Patson, 2007; see also Sanford & Sturt, 2002). These models assume that native language processing concurrently proceeds via two routes, on the one hand, rich and complete processing that encompasses all available syntactic, semantic, etc. information, and, on the other hand, less elaborate heuristic processing yielding incomplete but 'good enough' (Ferreira, 2003) or 'pseudosyntactic' (Townsend & Bever, 2001) representations of the input.

The LAST model holds that every sentence is comprehended twice. First, a shallow pseudosyntactic parse is computed on the basis of basic phrase-level segmentation, including assignment of syntactic categories, likely thematic roles, moving *wh*-fillers back into the gap positions, semantic information and frequency expectations. This incremental pseudoparse is immediately mapped to semantic representation. Second, a complete analysis yielding a fully specified syntactic representations is computed and checked against the semantic output of the pseudoparse.

In a similar vein, the model of 'good enough' processing claims that language comprehension involves a mix of heuristics and detailed syntactic analysis. 'Good enough' processing marks an initial assignment of interpretation to sentences based on canonical thematic templates (e.g. noun-verb-noun or agent-action-patient, Ferreira, 2003), animacy, plausibility and pragmatic inferencing, yet disregarding or deprioritizing morphosyntactic information. Several experiments show that comprehenders map an incomplete parse to interpretation that lingers and coexists with subsequent full syntactic (re)analysis. For instance, a range of experiments by Christianson and colleagues (Christianson, Hollingworth, Halliwell & Ferreira, 2001; Christianson & Slattery, 2005; Christianson, Williams, Zacks & Ferreira, 2006) investigates garden-path sentences such as in (10), in which the subject of the main clause is erroneously first analysed as the object of the verb in the embedded clause.

- (10) While Anna bathed the baby played in the crib.

In self-paced reading tasks, native readers are found to answer ‘yes’ to both the questions *Did Anna bathe the baby?* and *Did the baby play in the crib?* after reading the sentences. The affirmative responses to the second question attest that comprehenders correctly (re)analyse *the baby* as the subject of the matrix clause; however, the coexisting acceptance of the initial misinterpretation based on a ‘good enough’ interpretation of the sentence is not fully erased or overwritten by syntactic reanalysis. Ferreira and colleagues conclude that ‘good enough’ interpretations are the results of incomplete reanalysis, such that a string is mapped to interpretation before all sources of information are available for integration. Indeed, eyetracking experiments on sentences such as in (10) reveal that (a) the ambiguity is noted in incremental processing and (b) full reanalysis to the correct interpretation occurs in later processing even after the end of the sentence (Christianson & Slattery, 2005). Moreover, the degree of incomplete ‘good enough’ processing in sentences such as (10) is correlated with working memory (Christianson et al., 2006).

In light of these patterns, Ferreira (2003) argues that ‘good enough’ processing heuristics apply in part as a consequence of the dynamic demands of language processing in real time, noisy environments or under task pressure. Time constraints or processing pressures cut short the full parsing route or comprehenders economize the commitment of processing resources by relying on ‘good enough’ processing in communicative contexts that do not necessitate full parsing (see also Ferreira & Patson, 2007).

In sum, dual-route models of native sentence comprehension such as LAST or ‘good enough’ processing define incomplete processing as the immediate construction of sentence interpretation based on surface cues, associatively acquired patterns, lexical-thematic and semantic information when exogenous computational demands or endogenous computational limitations curtail use of the full parsing route.

How do these models apply to L2 processing? First, by positing two processing routes in sentence processing, dual-route models can capture potential differences between native and non-native processing in terms of computational load. On the account that the application of ‘good enough’ versus full processing is partly a function of computational load, incomplete processing in the L2 would occur sooner and in a more pronounced fashion than in monolingual processing. The fact that L2 processing is generally slower than native language processing may index higher load, possibly deriving from the parallel activation and inhibition of the L1, larger strains on working memory or reduced automaticity in L2 processing. Target-like performance in L2 processing should thus largely be a function of task demands, working memory capacity, proficiency level and L1 effects. Second, dual-route models allow for capturing differences between native and non-native processing without invoking fundamental differences in underlying grammatical architecture or mental processes, since both routes coexist in native and non-native language processing.

In sum, L2 capacity models posit that the processing architecture and mechanisms qualitatively conform to native processing, that is, unlike the two previous models, they do not claim fundamental differences in either language representation or processing between L1 and L2 processing. In addition, they integrate performance factors and capacity limitations in considering L2 processing. When capacities are taxed, processing heuristics such as incomplete ‘good-enough’ processing emerge. In L2 processing, capacity can be defined as a product of, e.g., (a) individual resources in working memory, reading speed, (b) computational complexity of particular linguistic structures or phenomena, (c) interference of L1 properties in L2 processing, and (d) exogenous task demands. This multifactorial compound renders it difficult to gauge the contributions of individual factors and extract conditions for (non-)convergence in the L2. Hence, capacity models only allow for limited predictions about convergence, such that a capacity explanation often remains post hoc.

Moreover, computational load is not defined relative to grammatical and processing architecture, such that it is not clear whether and, if so, which particular linguistic (sub-) domains are susceptible to capacity restrictions. In order to provide a substantive account of differences between native and non-native processing, computational capacity needs to be defined in a model of grammatical architecture and computational resources (see also Caplan et al. 2007). It will be interesting to study whether different grammatical domains or interfaces are differentially affected by capacity restrictions.

2.7.4. Summary of L2 processing models

In sum, the three models of (L2) processing reviewed above all accommodate the empirical evidence that L2 processing, at least at intermediate to advanced stages of proficiency, is characterized by greater reliance on surface, lexical-thematic and semantico-pragmatic information and comparatively shallow(er) syntactic analysis. However, the models postulate different aetiologies of divergent processing and different potentials for convergence. The ‘Shallow Structure’ Hypothesis posits that L2ers have qualitatively different grammatical representations, the DP model invokes different mental processing of complex syntax in processing, and capacity models emphasize computational limitations. Table 2.2 summarizes the predictions of the models regarding the main differences between native and non-native language processing. Cells marked with ‘?’ mean that the model in question does not assign a major role to the given factor, even though it may acknowledge its impact in L2 processing.

	Grammatical representations (L1 versus L2)	Processing architecture and mechanisms (L1 versus L2)	Computational resources (L1 versus L2)
Shallow Structure Hypothesis	different	same	?
Declarative/Procedural model	?	different	?
Capacity models	same	same	different

Table 2.2. *Properties of L2 processing models in terms of qualitative differences between L1 and L2 processing.*

In relation to the approaches to L2 acquisition discussed in Chapter 1, the ‘Shallow Structure’ Hypothesis resonates with approaches positing grammatical impairment in morphosyntax (the ‘Failed Functional Features’ Hypothesis Hawkins & Chan, 1997), although it argues for more extensive syntactic impairment affecting phrase structure. The capacity approach ties in with the ‘Missing Surface Inflection’ Hypothesis (Prévost & White, 2000b) in that they both highlight computational limitations in accessing and applying grammatical knowledge.

2.8. Summary and desiderata

We are now in a position to pull together previous research on non-convergence at L2 ultimate attainment in grammar acquisition (Chapter 1) and in processing (Chapter 2). This review has yielded a complex pattern of potentially interacting factors that might cause protracted non-native-likeness in adult L2 performance. By way of illustration, Figure 2.2, adapted from Figure 1.3 in Chapter 1, summarizes the models and situates them in the matrix of causes of non-convergence, its domain and its nature.

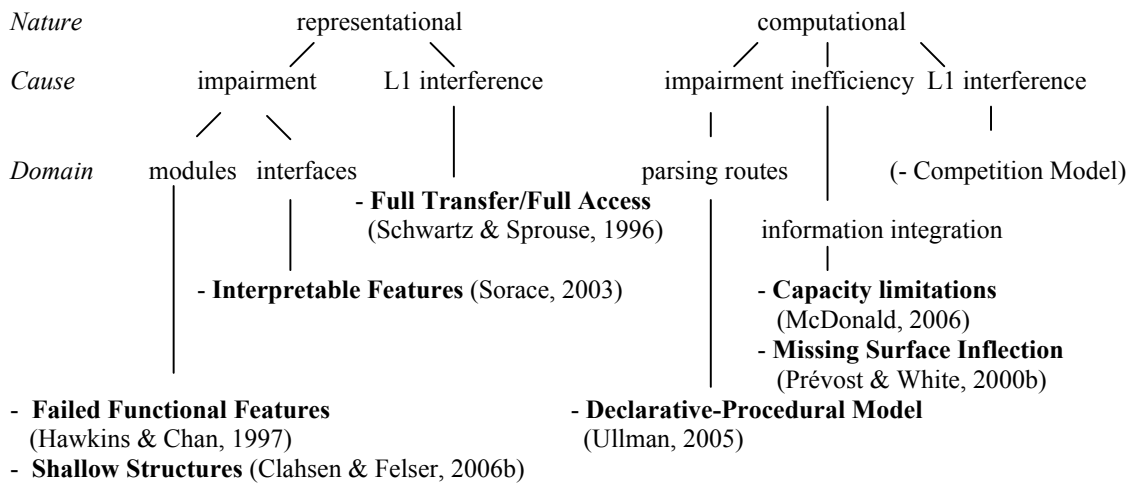


Figure 2.2. Approaches to non-convergence in adult L2 acquisition.

For the purpose of the experiments in this thesis, I group these models into three families of approaches. These approaches span research previously conducted in separate grammatical and processing perspective: (1) Grammatical impairment, (2) Different mental processing and (3) No impairment. (1) Assuming a critical period, grammatical impairment approaches locate representational differences between non-native and native grammars at the syntax-morphology interface of uninterpretable features (Hawkins & Chan, 1997; Tsimpli, 2003), the syntax-discourse interface of interpretable features (Sorace, 2003) or the syntax-parsing interface (Clahsen & Felser, 2006b). (2) Different mental processing holds that L2 speakers use partially different systems for processing the L2 (Ullman, 2005), partly as the consequence of a critical period. Finally, (3) assuming no critical period, no impairment approaches claim that non-convergence at L2 ultimate attainment results from performance limitations in accessing or integrating grammatical knowledge across interfaces in real time (e.g. McDonald, 2006; Prévost & White, 2000b) and/or persistent L1 transfer (Schwartz & Sprouse, 1996).

To date, there are not sufficient empirical data available to test between these approaches with respect to endstate L2 acquisition and, in particular, to ascertain how representational and computational factors might interact in giving rise to non-convergence at L2 ultimate attainment. This thesis aims to contribute some data in these respects. In Chapter 5, I will discuss the predictions these different approaches make for the present study.

2.8.1. Desiderata

The review of previous research on L2 grammars at ultimate attainment (Chapter 1) and L2 processing (Chapter 2) also identified gaps and shortcomings in our understanding of the factors potentially constraining convergence. In conclusion, I outline a number of desiderata for a study on L2 ultimate attainment that inform the design of the present study.

(1) Investigating different, yet related phenomena

Previous L2 research has attested non-convergence at ultimate attainment in various syntax-related domains. In order to be able to identify these domains of non-convergence in adult L2 acquisition, it is desirable for a study to test grammatical phenomena that comprise different aspects of syntax and its interfaces. These phenomena should be related, so that it can be determined which aspects are potentially subject to non-convergence at ultimate attainment. To this end, a phenomenon encompassing several syntactic and syntax-related aspects, word order optionality in German, namely, scrambling, will be the topic of investigation. In Chapter 3, a theoretical model of scrambling will be outlined and related to the psycholinguistic evidence available so far.

(2) Testing different L1 groups

In order to elucidate whether (non-)convergence on the TL depends on particular L1-TL pairings, L2 speakers of different L1 backgrounds will be tested. The choice of L1 groups will be motivated by specific cross-linguistic grammatical differences that lead to different conditions for transferring L1 properties and entail different learnability tasks.

(3) Testing L2 learners at different proficiency levels

In light of the findings of L2 processing studies showing that proficiency level underlies qualitative and quantitative differences between L1 and L2 processing, the present study will test L2 acquirers at different levels of advanced proficiency in order to consider effects of proficiency on convergence. Further, as, e.g., Birdsong (1992; 1999a; 2004) argues, limiting ultimate attainment studies to L2ers who were pre-screened for native-likeness restricts the findings to a sample of extraordinary L2ers. He suggests testing a sample of advanced L2 speakers presumed to be at the asymptote of L2 acquisition in order to (a) unearth potential factors that are prerequisites for convergence and (b) generalize the findings to L2 acquisition beyond sub-samples of extraordinarily successful L2 learners.

(4) Using off- and on-line methods

The combination of off-line and on-line methods in the study of L2 ultimate attainment can give more precise insight into the nature of non-convergence and address the question as to whether non-convergence is due to representational and/or computational differences between native and non-native mature systems. To this end, the present study will employ a conjunction of off-line and on-line tests with a particular emphasis on on-line data in order to explore the real-time comprehension of the TL. Investigating on-line comprehension at ultimate attainment can further address the different predictions of L2 processing models that have as yet not been tested on endstate L2 learners.

